

Listing of Claims:

Claim 1 (previously presented): A method for detecting beats in a compression encoded audio bitstream, said method comprising the steps of:

- (a) determining a baseline beat position using modified discrete cosine transform (MDCT) coefficients obtained from the audio bitstream;
- (b) deriving from the audio bitstream a window-switching pattern for sub-band sampling windows used to generate the MDCT coefficients;
- (c) determining a window-switching beat position based on the derived window-switching pattern;
- (d) comparing said baseline beat position with said window-switching beat position; and
- (e) validating said window-switching beat position as a detected beat if a predetermined condition is satisfied.

Claim 2 (original): A method as in claim 1 further comprising the step of determining an inter-beat interval related to said baseline beat position.

Claim 3 (original): A method as in claim 2 further comprising the step of storing said window-switching beat position and said inter-beat interval for subsequent retrieval.

Claim 4 (original): A method as in claim 1 wherein said step of determining a baseline beat position comprises the step of determining at least one beat candidate and an inter-onset interval.

Claim 5 (original): A method as in claim 4 wherein said step of determining a baseline beat position further comprises the step of checking said at least one beat candidate for reliability using a predetermined confidence threshold value.

Claim 6 (original): A method as in claim 4 further comprising the step of converging two or more said beat candidates to a single beat candidate.

Claim 7 (original): A method as in claim 1 wherein said step of deriving baseline beat information from the audio bitstream comprises the step of deriving an energy value for at least one subband from the compression encoded audio bitstream.

Claim 8 (original): A method as in claim 7 wherein said subband comprises a member of the group consisting of a frequency interval from 0 to 459 Hz, a frequency interval from 460 to 918 Hz, a frequency interval from 919 to 1337 Hz, a frequency interval from 1.338 to 3.404 kHz, a frequency interval from 3.405 to 7.462 kHz, and a frequency interval from 7.463 to 22.05 kHz.

Claim 9 (original): A method as in claim 7 wherein said step of deriving a beat position comprises the step of identifying a maximum energy value within a search window.

Claim 10 (original): A method as in claim 7 wherein said step of deriving an energy value for at least one subband comprises the step of deriving an absolute energy value.

Claim 11 (original): A method as in claim 7 wherein said step of deriving an energy value for at least one subband comprises the step of deriving an element-to-mean energy value.

Claim 12 (original): A method as in claim 7 wherein said step of deriving an energy value for at least one subband comprises the step of deriving a differential energy value.

Claim 13 (currently amended): A beat detector suitable for placement into an audio device conforming to a compression-encoded audio transmission protocol, said beat detector comprising:

a modified discrete cosine transform coefficient extractor, for obtaining transform coefficients from an audio bitstream;

at least one band feature value analyzer for analyzing a feature value for a related band, the at least one band feature value analyzer receiving input from the modified discrete cosine transform coefficient extractor;
a confidence score calculator receiving input from the at least one band feature value analyzer, the confidence score calculator calculating a confidence score for beat candidates using stored values of previous inter-onset intervals; and
a converging and storage unit for combining two or more of said analyzed band feature values-beat candidates.

Claim 14 (original): The beat detector as in claim 13 wherein said feature value comprises a member of the group consisting of an absolute energy value, an element-to-mean energy value, and a differential energy value.

Claim 15 (original): The beat detector as in claim 14 further comprising an element-to-mean ratio threshold comparator.

Claim 16 (original): An audio encoder suitable for use with a compression-encoded audio transmission protocol, said audio encoder comprising:

a beat detector including
a modified discrete cosine transform coefficient extractor, for obtaining transform coefficients;
at least one band feature value analyzer for analyzing a feature value for a related band;
a confidence score calculator; and
means for including beat detection information as side information in audio transmission.

Claim 17 (original): An audio decoder suitable for use with a compression-encoded audio transmission protocol, said audio decoder comprising:

a beat detector for providing beat position information, said beat detector including

a modified discrete cosine transform coefficient extractor, for obtaining transform coefficients;
at least one band feature value analyzer for analyzing a feature value for a related band;
a confidence score calculator; and
error concealment means for concealing packet loss in audio transmission by utilizing said beat position to identify audio data for replacement of packet loss.

Claim 18 (previously presented): The method of claim 1, wherein step (a) comprises determining a baseline beat position prior to inverse modified discrete cosine transform (IMDCT) processing of the MDCT coefficients.

Claim 19 (previously presented): The method of claim 1, wherein the predetermined condition of step (e) comprises relative displacement of the window-switching and baseline beat positions by less than a predetermined amount.

Claim 20 (previously presented): The method of claim 1, wherein step (a) further comprises:

- i) obtaining the MDCT coefficients from a portion of the audio bitstream within a search window,
- ii) sorting the MDCT coefficients into a plurality of subband divisions,
- iii) identifying beat candidates within some or all of the subband divisions,
- iv) calculating a confidence score for beat candidates identified in step iii),
- v) calculating a converged confidence score from the confidence scores of step iv),
and
- vi) determining the baseline beat position within the search window based on the converged confidence score.

Claim 21 (previously presented): The method of claim 20, wherein step iii) includes identifying a full band beat candidate across all of the subband divisions.

Claim 22 (previously presented): The method of claim 21, wherein step iv) includes calculating a confidence score using the following formula:

$$R_i = \max_{k=1,2,3} \left[\frac{\text{median}(\overline{IOI})}{\text{median}(\overline{IOI}) + \left| \text{median}(\overline{IOI}) - \frac{(I_i - I_{\text{last_beat}})}{k} \right|} \right] * f(E_i), \text{ wherein}$$

i is equal to F, 1, ..., N, where 1 through N are indices of subband divisions and F is the index for the full band,

R_i is equal to the confidence score for index i ,

\overline{IOI} is a vector of intervals between previous beat candidates within the subband divisions,

k is set to 1 unless the current interval between beat candidates within a subband division is two or three times longer than a predicted value because of a missed candidate, and set to 2 or 3 otherwise,

I_i is a granule index of a current beat candidate,

$I_{\text{last_beat}}$ is a granule index of a previous beat, and

$f(E_i)$ equals 0 if the energy (E) of a candidate for index i is less than a threshold, and is 1 if the energy (E) of that candidate is greater than the threshold.

Claim 23 (previously presented): The method of claim 22, wherein step v) includes calculating a converged confidence score using the following formula:

$$R_{\text{confidence}} = \max \{R_F, R_1, \dots, R_N\}.$$

Claim 24 (previously presented): The method of claim 20, wherein the search window size is adaptive.

Claim 25 (previously presented): The method of claim 24, wherein the search window is sized according to the formula

$$\text{window_size_new} = 2 * \text{floor}\left(\frac{\text{median}(\overline{IOI})}{2}\right) + 1, \text{ wherein}$$

window_size_new is a new size of the search window, and
 \overline{IOI} is a vector of intervals between previous beat candidates within the subband divisions.

Claim 26 (previously presented): The method of claim 20, wherein step iii) comprises identifying a feature value, within a subband division and during the search window, exceeding a threshold.

Claim 27 (previously presented): The method of claim 26, wherein identifying a feature value comprises determining whether a primitive band energy E within a subband division exceeds a threshold value, and wherein the primitive band energy E is calculated according to the formula

$$E_b(n) = \sum_{j=N1}^{N2} [X_j(n)]^2, \text{ wherein}$$

$E_b(n)$ is the energy of subband b in granule n,
 $X_j(n)$ is the jth normalized MDCT coefficient decoded at granule n,
N1 is a lower bound index of the MDCT coefficients sorted into subband b, and
N2 is an upper bound index of the MDCT coefficients sorted into subband b.

Claim 28 (previously presented): The method of claim 26, wherein identifying a feature value further comprises:

- (1) determining the energy in a granule,
- (2) determining the average energy in the search window,

(3) determining the ratio of the quantity determined in step (1) to the quantity determined in step (2).

Claim 29 (previously presented): The method of claim 26, wherein identifying a feature value further comprises computing a differential energy value for subband divisions using the formula $E_b(n+1) - E_b(n)$, wherein

$$E_b(n) = \sum_{j=N1}^{N2} [X_j(n)]^2 ,$$

$E_b(n)$ is the energy of subband b in granule n of the audio bitstream,

$X_j(n)$ is the jth normalized MDCT coefficient decoded at granule n,

N1 is a lower bound index of the MDCT coefficients sorted into subband b,

N2 is an upper bound index of the MDCT coefficients sorted into subband b,

$$E_b(n+1) = \sum_{j=N1}^{N2} [X_j(n+1)]^2 ,$$

$E_b(n+1)$ is the energy of subband b in granule n+1 of the audio bitstream,

$X_j(n+1)$ is the jth normalized MDCT coefficient decoded at granule n+1,

N1 is a lower bound index of the MDCT coefficients sorted into subband b, and

N2 is an upper bound index of the MDCT coefficients sorted into subband b.

Claim 30 (previously presented): The method of claim 1, wherein the audio bitstream is an MP3 encoded audio bitstream, and wherein step (b) comprises determining a pattern of long, long-to-short, short and short-to-long windows in the audio bitstream.

Claim 31 (previously presented): An audio encoder, comprising:
a beat detector, said beat detector being configured to perform a method for detecting beats in a compression encoded audio bitstream, said method including the steps of
(a) determining a baseline beat position using modified discrete cosine transform (MDCT) coefficients obtained from the audio bitstream,

- (b) deriving from the audio bitstream a window-switching pattern for sub-band sampling windows used to generate the MDCT coefficients,
- (c) determining a window-switching beat position based on the derived window-switching pattern,
- (d) comparing the baseline beat position with the window-switching beat position, and
- (e) validating the window-switching beat position as a detected beat if a predetermined condition is satisfied.

Claim 32 (previously presented): The audio encoder of claim 31, wherein step (a) comprises determining a baseline beat position prior to inverse modified discrete cosine transform (IMDCT) processing of the MDCT coefficients.

Claim 33 (previously presented): The audio encoder of claim 31, wherein the predetermined condition of step (e) comprises relative displacement of the window-switching and baseline beat positions by less than a predetermined amount.

Claim 34 (previously presented): The audio encoder of claim 31, wherein step (a) further comprises:

- i) obtaining the MDCT coefficients from a portion of the audio bitstream within a search window,
- ii) sorting the MDCT coefficients into a plurality of subband divisions,
- iii) identifying beat candidates within some or all of the subband divisions,
- iv) calculating a confidence score for beat candidates identified in step iii),
- v) calculating a converged confidence score from the confidence scores of step iv),
and
- vi) determining the baseline beat position within the search window based on the converged confidence score.

Claim 35 (previously presented): The audio encoder of claim 34, wherein step iii) includes identifying a full band beat candidate across all of the subband divisions.

Claim 36 (previously presented): The audio encoder of claim 35, wherein step iv) includes calculating a confidence score using the following formula:

$$R_i = \max_{k=1,2,3} \left[\frac{\text{median}(\overline{IOI})}{\text{median}(\overline{IOI}) + \left| \text{median}(\overline{IOI}) - \frac{(I_i - I_{\text{last_beat}})}{k} \right|} \right] * f(E_i), \text{ wherein}$$

i is equal to F, 1, ..., N, where 1 through N are indices of subband divisions and F is the index for the full band,

R_i is equal to the confidence score for index i ,

\overline{IOI} is a vector of intervals between previous beat candidates within the subband divisions,

k is set to 1 unless the current interval between beat candidates within a subband division is two or three times longer than a predicted value because of a missed candidate, and set to 2 or 3 otherwise,

I_i is a granule index of a current beat candidate,

$I_{\text{last_beat}}$ is a granule index of a previous beat, and

$f(E_i)$ equals 0 if the energy (E) of a candidate for index i is less than a threshold, and is 1 if the energy (E) of that candidate is greater than the threshold.

Claim 37 (previously presented): The audio encoder of claim 36, wherein step v) includes calculating a converged confidence score using the following formula:

$$R_{\text{confidence}} = \max \{R_F, R_1, \dots, R_N\}.$$

Claim 38 (previously presented): The audio encoder of claim 34, wherein the search window size is adaptive.

Claim 39 (previously presented): The audio encoder of claim 38, wherein the search window is sized according to the formula

$$\text{window_size_new} = 2 * \text{floor}\left(\frac{\text{median}(\overline{IOI})}{2}\right) + 1, \text{ wherein}$$

window_size_new is a new size of the search window, and

\overline{IOI} is a vector of intervals between previous beat candidates within the subband divisions.

Claim 40 (previously presented): The audio encoder of claim 34, wherein step iii) comprises identifying a feature value, within a subband division and during the search window, exceeding a threshold.

Claim 41 (previously presented): The audio encoder of claim 40, wherein identifying a feature value comprises determining whether a primitive band energy E within a subband division exceeds a threshold value, and wherein the primitive band energy E is calculated according to the formula

$$E_b(n) = \sum_{j=N1}^{N2} [X_j(n)]^2, \text{ wherein}$$

$E_b(n)$ is the energy of subband b in granule n,

$X_j(n)$ is the jth normalized MDCT coefficient decoded at granule n,

N1 is a lower bound index of the MDCT coefficients sorted into subband b, and

N2 is an upper bound index of the MDCT coefficients sorted into subband b.

Claim 42 (previously presented): The audio decoder of claim 40, wherein identifying a feature value further comprises:

- (1) determining the energy in a granule,
- (2) determining the average energy in the search window,

(3) determining the ratio of the quantity determined in step (1) to the quantity determined in step (2).

Claim 43 (previously presented): The audio decoder of claim 40, wherein identifying a feature value further comprises computing a differential energy value for subband divisions using the formula $E_b(n+1) - E_b(n)$, wherein

$$E_b(n) = \sum_{j=N1}^{N2} [X_j(n)]^2 ,$$

$E_b(n)$ is the energy of subband b in granule n of the audio bitstream,

$X_j(n)$ is the jth normalized MDCT coefficient decoded at granule n,

N1 is a lower bound index of the MDCT coefficients sorted into subband b,

N2 is an upper bound index of the MDCT coefficients sorted into subband b,

$$E_b(n+1) = \sum_{j=N1}^{N2} [X_j(n+1)]^2 ,$$

$E_b(n+1)$ is the energy of subband b in granule n+1 of the audio bitstream,

$X_j(n+1)$ is the jth normalized MDCT coefficient decoded at granule n+1,

N1 is a lower bound index of the MDCT coefficients sorted into subband b, and

N2 is an upper bound index of the MDCT coefficients sorted into subband b.

Claim 44 (previously presented): The audio decoder of claim 31, wherein the audio bitstream is an MP3 encoded audio bitstream, and wherein step (b) comprises determining a pattern of long, long-to-short, short and short-to-long windows in the audio bitstream.

Claim 45 (previously presented): An audio decoder, comprising:
a beat detector, said beat detector being configured to perform a method for detecting beats in a compression encoded audio bitstream, said method including the steps of
(a) determining a baseline beat position using modified discrete cosine transform (MDCT) coefficients obtained from the audio bitstream,

- (b) deriving from the audio bitstream a window-switching pattern for sub-band sampling windows used to generate the MDCT coefficients,
- (c) determining a window-switching beat position based on the derived window-switching pattern,
- (d) comparing the baseline beat position with the window-switching beat position, and
- (e) validating the window-switching beat position as a detected beat if a predetermined condition is satisfied.

Claim 46 (previously presented): The audio decoder of claim 45, wherein step (a) comprises determining a baseline beat position prior to inverse modified discrete cosine transform (IMDCT) processing of the MDCT coefficients.

Claim 47 (previously presented): The audio decoder of claim 45, wherein the predetermined condition of step (e) comprises relative displacement of the window-switching and baseline beat positions by less than a predetermined amount.

Claim 48 (previously presented): The audio decoder of claim 45, wherein step (a) further comprises:

- i) obtaining the MDCT coefficients from a portion of the audio bitstream within a search window,
- ii) sorting the MDCT coefficients into a plurality of subband divisions,
- iii) identifying beat candidates within some or all of the subband divisions,
- iv) calculating a confidence score for beat candidates identified in step iii),
- v) calculating a converged confidence score from the confidence scores of step iv),
and
- vi) determining the baseline beat position within the search window based on the converged confidence score.

Claim 49 (previously presented): The audio decoder of claim 48, wherein step iii) includes identifying a full band beat candidate across all of the subband divisions.

Claim 50 (previously presented): The audio decoder of claim 49, wherein step iv) includes calculating a confidence score using the following formula:

$$R_i = \max_{k=1,2,3} \left[\frac{\text{median}(\overline{IOI})}{\text{median}(\overline{IOI}) + \left| \text{median}(\overline{IOI}) - \frac{(I_i - I_{\text{last_beat}})}{k} \right|} \right] * f(E_i), \text{ wherein}$$

i is equal to F, 1, ..., N, where 1 through N are indices of subband divisions and F is the index for the full band,

R_i is equal to the confidence score for index i ,

\overline{IOI} is a vector of intervals between previous beat candidates within the subband divisions,

k is set to 1 unless the current interval between beat candidates within a subband division is two or three times longer than a predicted value because of a missed candidate, and set to 2 or 3 otherwise,

I_i is a granule index of a current beat candidate,

$I_{\text{last_beat}}$ is a granule index of a previous beat, and

$f(E_i)$ equals 0 if the energy (E) of a candidate for index i is less than a threshold, and is 1 if the energy (E) of that candidate is greater than the threshold.

Claim 51 (previously presented): The audio decoder of claim 50, wherein step v) includes calculating a converged confidence score using the following formula:

$$R_{\text{confidence}} = \max \{R_F, R_1, \dots, R_N\}.$$

Claim 52 (previously presented): The audio decoder of claim 48, wherein the search window size is adaptive.

Claim 53 (previously presented): The audio decoder of claim 52, wherein the search window is sized according to the formula

$$\text{window_size_new} = 2 * \text{floor}\left(\frac{\text{median}(\overline{IOI})}{2}\right) + 1, \text{ wherein}$$

window_size_new is a new size of the search window, and

\overline{IOI} is a vector of intervals between previous beat candidates within the subband divisions.

Claim 54 (previously presented): The audio decoder of claim 48, wherein step iii) comprises identifying a feature value, within a subband division and during the search window, exceeding a threshold.

Claim 55 (previously presented): The audio decoder of claim 54, wherein identifying a feature value comprises determining whether a primitive band energy E within a subband division exceeds a threshold value, and wherein the primitive band energy E is calculated according to the formula

$$E_b(n) = \sum_{j=N1}^{N2} [X_j(n)]^2, \text{ wherein}$$

$E_b(n)$ is the energy of subband b in granule n,

$X_j(n)$ is the jth normalized MDCT coefficient decoded at granule n,

N1 is a lower bound index of the MDCT coefficients sorted into subband b, and

N2 is an upper bound index of the MDCT coefficients sorted into subband b.

Claim 56 (previously presented): The audio decoder of claim 54, wherein identifying a feature value further comprises:

- (1) determining the energy in a granule,
- (2) determining the average energy in the search window,

(3) determining the ratio of the quantity determined in step (1) to the quantity determined in step (2).

Claim 57 (previously presented): The audio decoder of claim 54, wherein identifying a feature value further comprises computing a differential energy value for subband divisions using the formula $E_b(n+1) - E_b(n)$, wherein

$$E_b(n) = \sum_{j=N1}^{N2} [X_j(n)]^2 ,$$

$E_b(n)$ is the energy of subband b in granule n of the audio bitstream,

$X_j(n)$ is the jth normalized MDCT coefficient decoded at granule n,

N1 is a lower bound index of the MDCT coefficients sorted into subband b,

N2 is an upper bound index of the MDCT coefficients sorted into subband b,

$$E_b(n+1) = \sum_{j=N1}^{N2} [X_j(n+1)]^2 ,$$

$E_b(n+1)$ is the energy of subband b in granule n+1 of the audio bitstream,

$X_j(n+1)$ is the jth normalized MDCT coefficient decoded at granule n+1,

N1 is a lower bound index of the MDCT coefficients sorted into subband b, and

N2 is an upper bound index of the MDCT coefficients sorted into subband b.

Claim 58 (previously presented): The audio decoder of claim 45, wherein the audio bitstream is an MP3 encoded audio bitstream, and wherein step (b) comprises determining a pattern of long, long-to-short, short and short-to-long windows in the audio bitstream.